Information Technology Division

RESEARCH NOTE
ERL-0635-RN

DISTRIBUTED DATABASE MANAGEMENT EVALUATION SOFTWARE

by

S.J. Miller

APPROVED FOR PUBLIC RELEASE

AR-006-977

ELECTRONICS RESEARCH LABORATORY

Information Technology Division

RESEARCH NOTE
ERL- 0635-RN

DISTRIBUTED DATABASE MANAGEMENT EVALUATION SOFTWARE

by

S.J. Miller

SUMMARY

This paper describes techniques for assessing the performance of Distributed DataBase Manager (DDBN) systems. Both qualitative and quantitative methods of assessment are addressed. The desirability of collecting data from a test-bed system for use in the evaluation is highlighted. Finally, a functional requirement for a Distributed DataBase Management Evaluation Software (DBMES) package to be developed for the Distributed Processing Test-bed System (DPTBS) is presented.

AUG 92

APPROVED FOR PUBLIC RELEASE

POSTAL ADDRESS: Director, Electronics Research Laboratory, PO Box 1500, Salisbury, South Australia, 5108.

CONTENTS

Page No.

1 INTRODUCTION 1

2 ASSESSMENT OF DISTRIBUTED DATABASE SYSTEM PERFORMANCE 1

3 DISTRIBUTED DATABASE MANAGEMENT EVALUATION SOFTWARE 3
- 3.1 Distributed database 3
- 3.2 Communication interface 3
- 3.3 Exerciser 4
- 3.4 Sampler 4
- 3.5 Operator interface 4
  - 3.5.1 Input data entry mode 5
  - 3.5.2 Run mode 5
  - 3.5.3 DISPLAY task 5
  - 3.5.4 DDBM module 5

4 CONCLUSION 5

REFERENCES 7

Figure 1 Distributed Database Evaluation Software 9

1 INTRODUCTION

To fully utilise the added benefits of increased flexibility, reliability and availability, obtainable from distributed processing systems, the data used by the application programs executing on processors within the network must also be distributed and replicated over the processing nodes. Managing a database of this type in a manner that ensures the data remains consistent over the system at all times, including in the eventuality of network partitions and re-combinations, is a complex procedure. Control mechanisms are required for transaction concurrency control and dynamic re-configuration of the system to ensure that the data in the database replications converges into agreement and remains available to the applications. There are many such control protocols discussed in the literature(1). Concurrency mechanisms provide different degrees of control to meet varying transaction requirements. A means of comparing these different distributed database control mechanisms and hence evaluating the performance of different distributed database architectures and managers is required.

It is proposed that an evaluation software package be developed to run on the Distributed Processing Test-Bed System (DPTBS)(2), developed in Information Technology Division under the distributed processing task Nav 87/226.2, for evaluating database control mechanisms and concurrency protocols. This package should be modular in design and allow different database managers and transaction control mechanisms to be plugged in with minimal impact on it's basic evaluation functions.

In this paper, means of assessing and evaluating database systems are addressed. Section 2 discusses distributed database performance parameters and assessment procedures, and Section 3 specifies the functional requirement for a distributed database evaluation software package for the DPTBS.

2 ASSESSMENT OF DISTRIBUTED DATABASE SYSTEM PERFORMANCE

The assessment of the performance of different distributed database management systems is a complex procedure. It is first necessary to determine which parameters are of interest. In many situations a comparison of database concurrency control protocols in terms of their efficiency, transaction response time and transaction through-put, and their ability to maintain consistency over a network is sufficient. Where system availability and reliability are important, the robustness of these protocols in the event of network partitions is also important. A feature of importance in future Naval single-platform combat systems will be the system's ability to dynamically re-configure in cases of node damage and network partitions, the partitions each able to continue independent operation. On repair, the sections will need to be re-united as one and the distributed database replications need to remain consistent and converge to agreement. Other features of distributed databases include data independence, its transparency to the application, and its transparency to its distribution.

Parameters associated with the design of reliable and highly available distributed databases have been listed(3), under the headings of Base parameters, Control parameters, Failure parameters

and Performance parameters. An example of the type of parameters the author of the above article includes under these headings is shown below:

a. Base parameters

   (1) size of database

   (2) bandwidth and message delay of the network

   (3) number of copies

   (4) transaction read/update ratio and

   (5) node performance

b. Control parameters

   (1) type of copy

   (2) type of control

   (3) consistency

   (4) concurrency control mechanism

   (5) commit strategy

c. Failure parameters

   (1) failure model

   (2) availability

   (3) node recovery strategy

d. Performance parameters

   (1) transaction response time

   (2) transaction throughput

Garcia-Molina also suggests that the Performance Parameters of transaction response time and through-put can be best improved by trade-offs in the other parameters.

Much of the database literature deals with the comparison of Update algorithms very much in isolation from the other requirements of Distributed Databases when determining performance information. In one such article(4), algorithms are compared at two levels:

- a qualitative level : the algorithms are classified according to the kind of concurrency control mechanism used in maintaining consistency;

- a quantitative level : useful parameters are listed and results such as response time, number of needed messages, are computed.

At the quantitative level, techniques such as analysis models and simulation models can be very useful. In addition experimental techniques using instrumented test-beds are an appropriate way of getting real and accurate results suitable for validating models. It is suggested that a thorough evaluation of a distributed database management system needs to address all the above considerations.

3 DISTRIBUTED DATABASE MANAGEMENT EVALUATION SOFTWARE

The proposed special purpose Distributed DataBase Management (DDBM) evaluation software package is to be referred to as the Distributed dataBase Management Evaluation Software (DBMES). DBMES is to be hosted on the DPTBS to provide an experimental test-bed system for exercising different database control protocols and obtaining real data to validate simulation models. It is to be a modular package incorporating tasks for, operator control (OPERATOR INTERFACE), on-line display (DISPLAY), exercising the database control and concurrency protocols being evaluated (EXERCISER), inter-node communication (COMMUNICATION INTERFACE), and sampling the state of the database being managed (SAMPLER). To evaluate different DDBMs, the DBMES (figure 1) must be reconfigurable, requiring only minor modifications, mainly to the OPERATOR INTERFACE, the EXERCISER pre-programmed transaction sequences, and the DISPLAY task. The whole software package, together with the special DDBM module, its accompanying communication interface module (COMMUNICATION INTERFACE), and the database under test, must be compiled and linked to form one executable program.

3.1 Distributed database

To allow easier comparison, a standard distributed database is required for use with each DDBM system under test. The database must be suitable for the evaluation of fully replicated and partially replicated database systems. The granuality of replication is the page and the granuality of lock is the record. There should be headers, containing control information, attached to all pages and records. These headers should contain parameters, such as update number, time of update, frequency, and access type, appropriate to the protocol under test.

3.2 Communication interface

The COMMUNICATION INTERFACE module is to essentially provide the link between the DDBM and the Fiber Distributed Data Interface (FDDI)(2) which is the communication medium between the processing nodes. The DDBM module will initiate database access requests which must be passed to other processing nodes. The COMMUNICATION INTERFACE should format these requests into database control messages for transmission on the FDDI. It must also pass back any subsequent reply messages to the relevant DDBM module function. Because of its close links with the DDBM, the Communication Interface must be considered as part of it and each time a new DDBM is developed the Communication Interface will probably need to be changed.

Additional functionality must be built into the Communication Interface for the DBMES to allow it to format data records and store them in a communication data file for periodic dumping to disk storage for off-line evaluation (see figure 1). These records should contain information relating to type, source and destination of read and update requests,

and subsequent replies. The relevant data is to be extracted from the communication I/O messages. Other data records containing DDBM interface related data (see Section 3.6), must also be stored in this data file. The data file, together with a snapshot of the local database, is to be dumped to disk storage by the SAMPLER function. The sample period will be set by an operator at only one node. This node is to be referred to as the master (probably node 1). The snapshots taken at each node must be synchronised so that the data in the database can be checked for consistency.

3.3 Exerciser

An application task, to be known as the EXERCISER, is required to sequence through a pre-selected transaction sequence. It should have a set of pre-programmed transaction functions performing read, compute and updating of selected local database records (Section 3.1). The read and update function calls are to provide a standard interface to the DDBM. Transaction sequences are to be developed to suit the concurrency protocols under test (performance or reliable). An operator should be able to select independently at each node (through the OPERATOR INTERFACE) a pre-programmed sequence, the repetition period for the selection, and the total run time.

3.4 Sampler

The SAMPLER task at each node is to be initiated periodically and in synchrony (Section 3.2) at the rate set by the operator at the master node. This synchronism is achieved by the master node sending a broadcast message to initiate the SAMPLER task at each node. Each time SAMPLER is initiated, a snapshot of the local database is taken and stored together with the communication data file and local time on hard disk for later off-line evaluation. The SAMPLER task must be given the highest priority so that when it is initiated, all other node processing is to be halted until the sampling is completed. When SAMPLER terminates, the task suspended by it must be resumed. In this way, operation of the DDBM is suspended while the SAMPLERs take their snapshots.

3.5 Operator interface

The operator should be able to control the DBMES package through the OPERATOR INTERFACE by means of the PC keyboard and monitor. Control data is to be input via the PC keyboard, and message data displayed on the PC monitor. In the data entry mode, the following data and selections are required to be entered by the operator:

a. the relevant Node_Number,

b. selection of the appropriate transaction sequence,

c. the required repetition rate,

d. setting of the run time for the sequence,

e. the monitor display format (normal or demonstration), and

f. selection of RUN to start the local node sequence.

An operator must be able to return to the data entry mode at any time by hitting "ESC".

3.5.1 Input data entry mode

In the Input Data Entry mode, the operator shall be presented with an input data entry page on the monitor. This page should provide a numbered list of items requiring parameter input.

3.5.2 Run mode

Selecting RUN initiates the EXERCISER task and the display of page data on the monitor. Typically, the Normal display page could provide the operator with a summary of the database access and inter-communication resulting from the local and external node transaction sequences. The data for display being derived periodically from the communication data file, the local database, and the FDDI input/output message rates from the COMMUNICATION INTERFACE.

3.5.3 DISPLAY task

The DISPLAY task should provide the operator with a view of the performance data. The parameters displayed on the monitor could be different for each type of database architecture and DDBM under test. Two display formats are to be provided, Normal and Demonstration. The required format is to be selected prior to initiating the RUN process (Section 3.5.2). The Normal display should be of detailed performance data relating to the DDBM and database under test while, the Demonstration display format is to be used to display a situation summary which provides a simplified view of the current database configuration and content.

3.5.4 DDBM module

The DDBM module provides the application, in this case the EXERCISER, with an interface to the standard database. This module will be replaced each time a new distributed update control protocol is to be evaluated. The DDBM must contain the required distributed database control algorithms associated with the update functions and must format the required inter-node database requests, send them to the co-operating COMMUNICATION INTERFACE, and act on the replies received back from the COMMUNICATION INTERFACE. Further, all DDBM modules must implement the standard user interface to provide commonality between themselves and the EXERCISER.

A special purpose function is to be included in the DDBM for extracting the relevant performance data and storing this data in the communication data file.

The DDBM should include a transaction manager incorporating the special purpose transaction control protocols and any database configuration manager required for the special case of a partition and replicated database. The fully replicated database does not require a database configuration manager.

4 CONCLUSION

This paper introduces methods for assessing the performance of distributed database manager control mechanisms. It is suggested that both quantitative and qualitative assessment methods

should be used. An experimental method using an instrumented test-bed is also a useful way of obtaining real data for validating simulation models.

The latter part of this document was devoted to the functional requirement of a software evaluation package, the DBMES, to be hosted on the DPTBS. The DBMES will allow real data to be obtained for simulation model validation.

The above techniques and the DBMES will be used to analyse the performance of transaction control mechanisms and evaluate DDBMs for their suitability to Naval single platform command and control systems.

REFERENCES

No.	Author	Title
1	Miller, S.J.	"Distributed Databases for Command and Control: Task Overview". WSRL-TM-46/88, July 1988
2	Miller, S.J.	"Distributed Processing Test-Bed System: First Interim Report for the DPTBS". WSRL-TM-26/90, September 1990
3	Garcia-Molina, H.	"The Future of Data Replication" IEEE Symposium Reliability in Distributed Software and Database Systems, 1986
4	Wilms, P.	"Qualitative and Quantitative Comparison of Update Algorithms in Distributed Databases". Distributed Databases, Delobel, C. and Litwin, W. (eds.), North-Holland Publish Company, 1980

THIS IS A BLANK PAGE

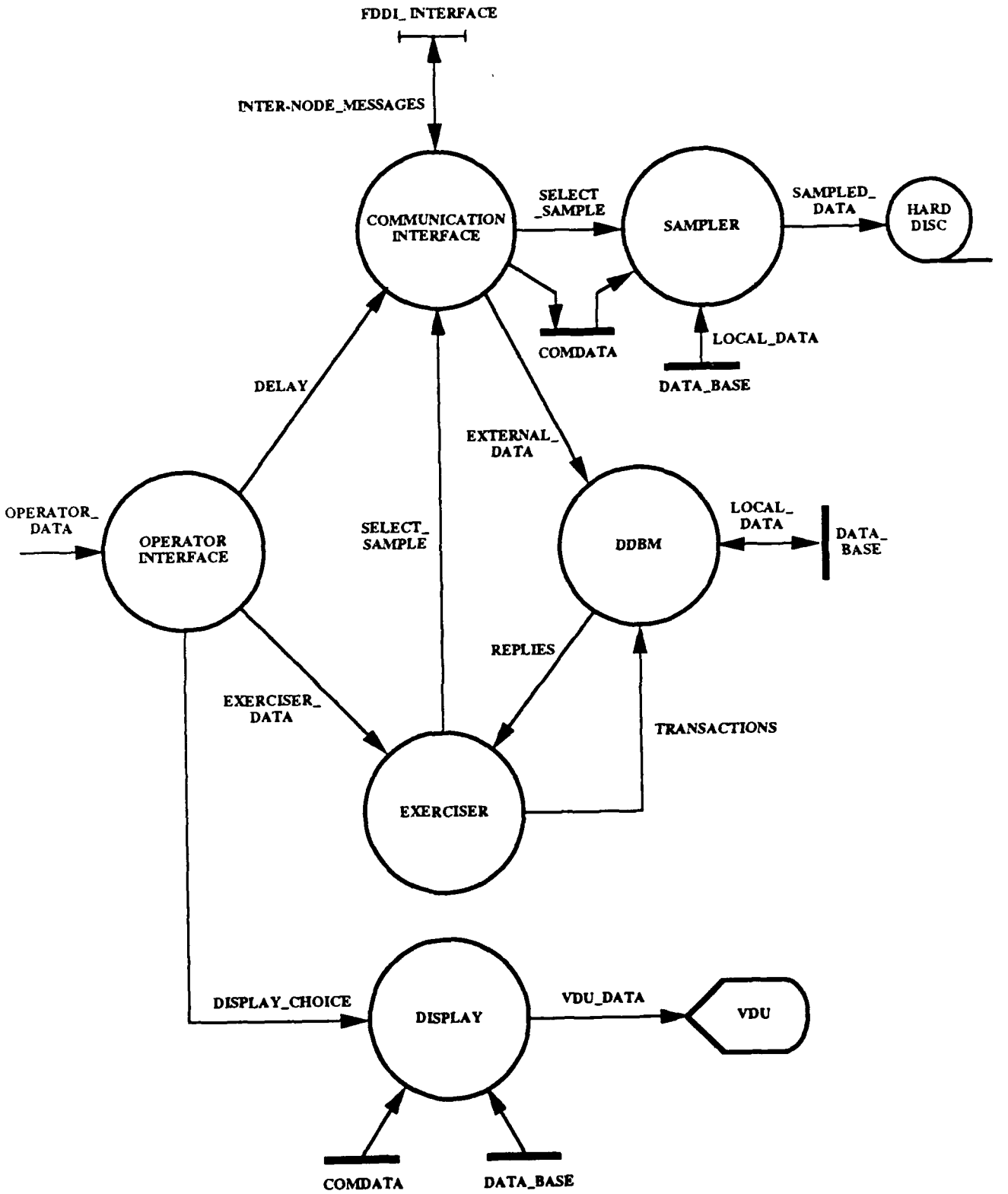

Figure 1 Distributed Database Evaluation Software

DISTRIBUTION

	Copy No.
Defence Science and Technology Organisation	
Chief Defence Scientist)	
Central Office Executive)	1 shared copy
Counsellor, Defence Science, London	Doc Cont Data Sht
Counsellor, Defence Science, Washington	Doc Cont Data Sht
Scientific Adviser, Defence Central	1
Scientific Adviser, Defence Intelligence Organisation	1
Mr P. Drewer, PRS Intelligence Development and Systems DIO	1
Director General, Communications and Information Systems	1
Navy Scientific Adviser	1
Director, Naval Combat Systems Engineering	1
Air Force Scientific Adviser	1
Scientific Adviser, Army	1
Electronics Research Laboratory	
Director	1
Chief, Information Technology Division	1
Chief, Electronic Warfare Division	1
Chief, Communications Division	1
Research Leader Command and Control and Intelligence Systems	1
Research Leader Military Computing Systems	1
Research Leader Human Computer Interaction	1
Head, C3I Systems Engineering Group	1
Head, Program and Executive Support ITD	1
Head, Trusted Computer Systems Group	1
Head, Software Engineering Group	1
Head, Computer Systems Architecture Group	1
Head, Command Support Systems Group	1
Head, Image Information Group	1
Head, Information Acquisition and Processing Group	1
Head, Information Management Group	1
Head, Exercise Analysis Group	1
Mr S. Miller, (C3ISE) Author	2
Mr J. Schapel, (C3ISE)	1
Mr A. Allwright, (C3ISE)	1
Mr D. O'Dea, (C3ISE)	1
Publications and Publicity Officer ITD	1
Media Services	1
Libraries and Information Services	
Australian Government Publishing Service	1
Defence Central Library, Technical Reports Centre	1

Manager, Document Exchange Centre, (for retention)	1
National Technical Information Service, United States	2
Defence Research Information Centre, United Kingdom	2
Director Scientific Information Services, Canada	1
Ministry of Defence, New Zealand	1
National Library of Australia	1
Defence Science and Technology Organisation Salisbury, Research Library	2
Library Defence Signals Directorate, Melbourne	1
British Library Document Supply Centre	1
Spares	
Defence Science and Technology Organisation Salisbury, Research Library	6

Department of Defence

UNCLASSIFIED

DOCUMENT CONTROL DATA SHEET

Privacy Marking/Caveat
(of document)

1a. AR Number	1b. Establishment Number	2. Document Date	3. Task Number
AR-006-977	ERL-0635-RN	AUG 92	

4. Title

DISTRIBUTED DATABASE MANAGEMENT EVALUATION SOFTWARE

5. Security Classification

Document	Title	Abstract
U	U	U

S (Secret) C (Confi) R (Rest) U (Unclass)

* For UNCLASSIFIED docs with a secondary distribution LIMITATION, use (L) in document box.

6. No. of Pages: 12

7. No. of Refs.: 4

8. Author(s)

S.J. Miller

9. Downgrading/Delimiting Instructions

N/A

10a. Corporate Author and Address

Electronics Research Laboratory
PO Box 1500
SALISBURY SA 5108

10b. Task Sponsor

DSTO

11. Officer/Position responsible for

Security.....

Downgrading..... N/A

Approval for Release..... DERL

12. Secondary Distribution of this Document

APPROVED FOR PUBLIC RELEASE

Any enquiries outside stated limitations should be referred through DSTIC, Defence Information Services, Department of Defence, Anzac Park West, Canberra, ACT 2600.

13a. Deliberate Announcement

No limitation

13b. Casual Announcement (for citation in other documents)

[√] No Limitation

[ ] Ref. by Author , Doc No. and date only.

14. DEFTEST Descriptors

Distributed database systems, Data base management systems, Distributed processing, Performance evaluation

15. DISCAT Subject Codes

1205

16. Abstract

This paper describes techniques for assessing the performance of Distributed DataBase Manager (DDBN) systems. Both qualitative and quantitative methods of assessment are addressed. The desirability of collecting data from a test-bed system for use in the evaluation is highlighted. Finally, a functional requirement for a Distributed DataBase Management Evaluation Software (DBMES) package to be developed for the Distributed Processing Test-bed System (DPTBS) is presented.

Doc Sect WF11

Page Classification

UNCLASSIFIED

16. Abstract (CONT.)

17. Imprint

Electronics Research Laboratory
PO Box 1500
SALISBURY SA 5108

18. Document Series and Number	19. Cost Code	20. Type of Report and Period Covered
ERL-0635-RN		RESEARCH NOTE

21. Computer Programs Used

N/A

22. Establishment File Reference(s)

N/A

23. Additional information (if required)